ONLINE SUBJECT SEARCHING OF DISSERTATIONS

Eva Bratková

brt@cuni.cz

Charles University, Institute of Information Studies and Librarianship, Czech Republic

Abstract

This paper evaluates searching for doctoral dissertations by subject in various online systems. The situation in the Czech Republic is introduced, including the problems involved in completing successful subject searches for dissertations – is it possible to find all relevant materials, or is it sufficient just to find something? The Czech situation is then compared with how systems abroad, particularly in the United States, are being implemented with new access routes for dissertations in the form of linked open data, in which controlled vocabularies of subject terms figure prominently. The paper also discusses how selected European systems whose dissertations are already presented in the WorldCat database will cope with a challenge: „... Over time, these references [for topic entities] will be replaced with persistent URIs to... Linked Data resources"?.

Keywords

Dissertations, Bibliographic Records, Metadata, Subject Control, Abstracts, Keywords, Controlled Vocabularies, Information Retrieval, Digital Repositories, Linked Open Data

Introduction

Doctoral dissertations are some of the **most important information outputs** from universities. They contain valuable content and include the results of scientific research. Dissertations are a source that should be – like scientific books – properly bibliographically processed, including the subjects. As American librarians also often state, it is very much worth preparing the widest and best possible access to dissertations, and so they prepare full catalog records of them, including subject terms, using controlled vocabularies (subject heading systems, classification schemes, etc.) (Middleton, 2015, p. 235; McCutcheon, 2008, p. 51). Well prepared dissertation records can then today be effectively linked through semantic web technologies in the form of open linked data. The linking of dissertation records with **subject authority records** is essential for today's users.

The bibliographic registration of dissertations at national or international level has a long tradition in many countries. Also typical for the new millennium is their effective storage in digital repositories to make them easily searchable online and, especially, accessible. The metadata that accompanies dissertations in local digital systems then also plays an important role in services at national and international level, into which they are easily harvested. The online searching for dissertations using formal bibliographic data (author, title, language, university, assigned degree name, etc.) is usually problem–free and, as a rule, also successful. There is, however, a problem with **searching for dissertations by subject**. If the search is dependent on an index composed solely of terms from the titles of the dissertations, from the authors' keywords and the abstract, or possibly also from the texts themselves, the **recall** of the search is usually low (while **precision** may also be high), as shown by the results of earlier well–known research[1]. Collections of dissertations and the records about them continue to grow, and searching by subject is hence becoming a more complicated problem, in particular from the perspective of professional searchers. The end user as a rule always finds "something" and then selects "something" from the offered list without realising that there are many more highly relevant dissertations in the repository.

For the reasons described above, this paper focuses exclusively on **dissertations** prepared within the framework of doctoral programmes. Other types of theses (bachelor's, master's, etc.) are not analysed or assessed, even though they also tend to be, in particular in low–output countries, registered in both local and national systems.

Searching for doctoral dissertations by subject in the Czech Republic

In the Czech Republic today, it is not possible to search for records of all defended dissertations from a single location, nor those defended in the past 10 years (2007 to 2016), meaning the period when they are already mandatorily publicly accessible everywhere. Using existing information services, it is not even possible to establish how many dissertations were actually successfully defended in individual years. The national services *Theses.cz* (**http://theses.cz/**) and *NRGL* (the *National Repository of Grey Literature*, **http://nusl.cz**) report approximately the same annual additions – around 1,000 defended dissertations. Yet

---

[1] BLAIR, David C., MARON, M. E. An evaluation of retrieval effectiveness for a full-text document-retrieval system. In: *Communications of the ACM*. 1985, **28**(3), 289-299. ISSN 0001-0782 (Print). ISSN 1557-7317 (Online). Available commercially from: **http://dx.doi.org/10.1145/3166.3197**. Freely available as a report from: **http://hdl.handle.net/2027.42/35415**.

the real figure is probably higher, as not all dissertations are registered at national level, especially from large universities. There could be over 2,000 defended dissertations a year.

Years of effort[2] to create complete national registration have still not been successful. The *NRGL* service has the highest number of dissertation records in its database: in a declared retrospective covering 1922 to 2016 it had – in September 2016 – a total of **18,304 records**, while for 2007 to 2016 it has **7,496**. The *Theses.cz* service has a slightly lower number of dissertation records in its database for the 2007 to 2016 period (**7,410** in September 2016), whereas the service has **7,716 records in total**. However, we can note that the two services do not carry out registration in the same way (see Table 1) – some dissertations are the same, some are different. Many dissertations are missing and can only be searched for in university databases. It is up to the individual universities as to which service they send their records to, or have them automatically harvested from the repositories. In the Czech Republic, there is currently no legislation for mandatory national registration, and voluntary reporting is not typical in this country.

Searching for dissertations by subject in the Theses.cz system

The national register of theses – *Theses.cz* – managed by Masaryk University, serves, inter alia, for searching for records of these works and the works themselves. The user interface has not changed greatly for some years, is not particularly user friendly either when formulating queries or when evaluating the resulting lists of records and the presentation of full records. It offers searches by several formal data types (year of defence, name of university, etc.). Searches by type of work is surprisingly missing, this being replaced by a view of assigned degree names (however, the results might not be accurate).

Figure 1: "Cloud" of alphabetically ordered common keywords in the Theses.cz system

[2] Efforts since 1990, intended to build upon the earlier complete bibliographic registration of dissertations in the Czech Republic and also in the Slovak Republic: *Bibliografický katalog ČSSR. České knihy. Zvláštní sešit. Československé disertace …* 1964-1978. Praha: Státní knihovna ČSSR, 1965-1981. ISSN 0323-1763. -- *Bibliografický katalog ČSSR. České knihy. Zvláštní sešit. České disertace …* 1979-1988. Praha: Státní knihovna ČSSR, 1983-1990. ISSN 0232-041X.

Searching for dissertations by subject however brings its own problems. The system basically offers a single option: search by **keywords** (simple terms or phrases). The user must enter their query into a field in a simple interface. An index of alphabetically ordered keywords is not available. Some compensation is the option to browse a small quantity of highly common keywords selected from the whole collection (however, all types of theses are included in this case). The set is presented in the form of an impressive "cloud" – see Figure 1 – however, it cannot be restricted to dissertations only. Clicking on a selected term produces an extensive list of resulting short records of all types of theses, in which it is difficult to identify and select highly relevant dissertations.

Ordinary users have perhaps over the years become used to a situation in which they enter a simple keyword the system uses to present a list of selected records, however definitely not all of which are relevant in terms of content. We can illustrate this using the example of the "digitální knihovny" subject (in Czech). Entering this query using various synonymous or variant terms (something an end user would however probably not usually do) led, in September 2016, to different results each time: "digitální knihovny" (466 records), "digitální knihovna" (up to 532), "elektronická knihovna" (184), "elektronické knihovny" (145), "digitální repozitář" (105), "digitální repozitáře" (20), "digitální archiv" (252), "digitální archivy" (27), "institucionální repozitář" (26) and so on. Assembling these together is difficult for a user – the manual processing of long lists of records is difficult, as is evaluating the level of content relevance for each record. An attempt to search for works on the subject of the actual *Theses.cz* system itself ("theses.cz") led, in September 2016, to 993 records, however most of these records were irrelevant.

As a bonus, the *Theses.cz* system offers searches for works on a "related topic" (related subject) in the form of a reference in searched short records. This is a computer calculation based on the presence of common keywords in records of works. The obtained set of records of related works is, however, also difficult to evaluate manually as regards content relevance – the user tends to get lost in the list, and if they click more than once, they are completely lost. It is difficult to discover highly relevant works.

The *Theses.cz* system, which is based, as regards subjects, only on author keywords and abstracts from university repositories or study systems, still has difficulty providing high quality and effective dissertation searches according to subject. Linking records to subject terms in any controlled vocabularies is not yet possible.

Searching for dissertations by subject in the NRGL system

The *NRGL* system is managed by the National Library of Technology (NTK). Up to 75% of the *NRGL* system is composed of records of all types of theses from Czech universities (Charvátová, 2016, p. 91). The central user interface (FAST software), built over the actual *NRGL* repository (Invenio software), is modern and user–friendly, and has text filters for comfortable navigation. As regards dissertations, this service also offers searching by several types of formal bibliographic data (author, name of university, language, type of work etc.) Limiting searches by year of defence is configurable using an impressive timeline. In addition to browsing and combining lists of data, it is also possible to directly enter queries into the system.

Searching for dissertations by subject is also in this case based only on author keywords, which are part of the records harvested from university repositories or databases of study systems of selected universities. The efficiency of the searches is again not optimal in this service. The terms representing the subjects are not controlled. A small check test performed by searching using the subject "digitální knihovny" (in Czech) by entering various synonymous or variant terms also led in this case to different results each time (in September 2016): "digitální knihovny" (346 records), "elektronická knihovna" (350), "digitální repozitář" (61), "digitální repozitáře" (22), "digitální archiv" (47) and "institucionální repozitář".

Efforts to improve the state of affairs in terms of the subject description and online searching for works by subject led a *NRGL* workplace to carry out a small survey of the state of author keywords with the objective of finding whether it would, for example, be possible to map them to the Polythematic Structured Subject Heading System (PSH)[3], which is managed and used for subject description at the NTK as a controlled vocabulary. In addition, it is also prepared in the form of linked open data. The results so far have shown that the mapping would be significantly ineffective (Charvátová, 2016, p. 91-95).

Unless the description of dissertations in source digital systems of universities – from which metadata is harvested for the NRGL – changes or improves, the effectiveness of searches by subject will not significantly change either. The question is whether it would be better to harvest bibliographic metadata from the catalog databases of universities instead, as these pay more attention to subject description in the majority of cases at professional level. What is preventing this?

Description of dissertation subjects in local systems of Czech universities

The unsatisfactory results of searches for Czech dissertations by subjects in both national services are caused by the fact that the source records for the dissertations come from digital repositories of universities or from databases of their study systems, while the description of the subjects was provided by the authors themselves (keywords and abstracts). This means that searches by subject in local repositories also show the same parameters described above for the national services. Repositories also offer users (in the DSpace application or elsewhere) under the label "Browse by subject" (or other label) sometimes longer lists of terms referring to the same thing – see the sample terms on the subject "informační systémy" (in Czech) from the otherwise great–looking repository of the Brno University of Technology (BUT).

---

[3] *Polytematický strukturovaný heslář (PSH)* [online]. Praha: Národní technická knihovna, 2000- [cit. 2016-09-26]. Available from (URI): **http://pshmanager.techlib.cz/**.

Figure 2: The index of "subject headings", or author keywords on the subject "informační systém" ("information system" in English) in the BUT repository (English equivalents are in a different part of the alphabet)

For the purposes of this paper, a small survey was performed of the **description of the subjects of dissertations** defended at selected Czech universities[4], including outside digital repositories. The objective was to establish if subject description is better elsewhere, if the searchability of dissertations by subject is better elsewhere, and whether there is a chance – in the near future – to get linked records about them in the context of the semantic web. As Table 1 shows, the description of dissertation subjects is performed in **institutional repositories (IR)** exclusively through author abstracts (AB) and keywords (KW). Yet both elements – or even one of them – are however not usually present in many records at all (the policies of the universities differ in this area). At two universities, repositories either do not exist or do not include dissertations (Czech University of Life Sciences (CULS) Prague and Palacky University (PU) Olomouc), or the records are only available in the databases of the study systems.

An interesting although disparate approach to author abstracts and keywords is, however, adopted by the **centralised catalogues of university libraries**[5]. 6 of the 14 catalogues take abstracts from the repositories, while up to 9 catalogues take keywords. Author keywords are, however, not localised uniformly into the MARC 21 format field, some catalogues are localised

---

[4] The 14 universities that have reported the highest numbers of defended dissertations in recent years. The numbers of dissertations were taken from both national services, and also from databases of local systems, including the central catalogues.

[5] Of the 14 universities, 3 do not catalogue dissertations at all, however rare records do sometimes appear (see Table 1).

into field 690, and the majority into field 653 (Index Term-Uncontrolled)[6]. The Masaryk University catalogue also uses the local field M530. Keywords are usually displayed in OPAC records, however are not, for understandable reasons, filed into the indexes of subject terms.

What was confirmed through the survey as pleasing and hopeful was that the majority of universities take great care when describing the subjects in their catalogues. Sometimes the processing is actually very detailed. 8 of the 14, respectively 11 universities, use one of the well–known domestic **controlled subject heading systems or thesauri** (see Table 1). This is primarily the Czech National Subject Authority File of the National Library of the Czech Republic (CZENAS), the Polythematic Structured Subject Heading System (PSH), and also the CZMESH subject heading system (for medicine). Other systems are also used (Czech theological thesaurus, CTT), and also local subject heading systems (CZ-BrMU, PHFFUK, MFF etc.) In the case of the PSH, there is the possibility of immediate linking of the catalog records to the PSH, which is already displayed in a structured form (the subject headings have URI identifiers), while with CZENAS there will probably be the possibility to display it in the form of linked open data in the near future, as can be anticipated from the conclusions of the INTERPI[7] research project. Up to 9 catalogues use the **UDC classification scheme** (Universal Decimal Classification) to describe dissertation subjects. In this case, there is definitely a possibility of an open linking of records to records of all UDC numbers. For now, there is a set of selected main UDC numbers available in a linked open data structure (**http://udcdata.info/**). For example, the class of medical sciences has a URI: **http://udcdata.info/037318**. Access to the whole UDC set (**http://cz.udc-hub.com/cs/login.php**) is available in Czech after registration. In the future, when record volumes will grow quickly, the orderliness of dissertations according to this universal classification will definitely be of benefit.

---

[6] In one case a curiosity was discovered: author keywords in catalog records of the University of Economics, Prague, are localised in field 690, yet in the NRGL database that collects these records, they are then offered for export in field 653.

[7] The headquarters of the INTERPI project is at the URL: **http://autority.nkp.cz/interpi**.

	IR		LIBRARIES CATALOGUE					THESES.CZ		NRGL		BASE			OpenAIRE		DART-EU	
	AB	KW	AB	KW	SH	UDC	conspectus	AB	KW	AB	KW	AB	KW	DDC	AB	KW	AB	KW
CULS, Prague			—	—	CZENAS	UDC	conspectus											
CTU, Prague	AB	KW	—	—	PSH	UDC	conspectus			*AB*	*KW*							
USB, Čes.Bud.			—	KW	CZENAS	UDC	conspectus	AB	KW	AB	KW				AB	KW		
MU, Brno	AB	KW	AB	KW	CZENAS CZ-BrMU CZMESH ET AL.	UDC	conspectus	AB	KW									
MENDELU	AB	KW	AB	KW	—	—	—	AB	KW	AB	KW				AB	KW		
TU Liberec	AB	KW	AB	KW	CZENAS	—	conspectus	AB	KW	*AB*	*KW*							
Charles Univ.	AB	KW	AB	—	CZENAS PHFFUK CTT, MFF CZMESH ET AL.	UDC	conspectus											
PU, Olomouc			AB	KW	—	UDC	—	AB	KW	*AB*	*KW*							
UPA	AB	KW								*AB*	*KW*	AB	KW	DDC				
TBU, Zlin	AB	KW	—	KW	CZENAS ECZENAS ET AL.	UDC	conspectus	AB	KW			AB	KW	—				
VŠB-TUO	AB	KW						AB	KW	*AB*	*KW*	AB	KW	DDC				
UE, Prague	AB	KW	AB	KW	CZENAS	UDC	—	AB	KW	AB	KW	AB	KW	DDC	AB	KW	AB	KW
BUT, Brno	AB	KW	—	KW	PSH	UDC	—			AB	KW	AB	KW	DDC	AB	KW		
UWB, Pilsen	AB	KW						AB	KW	*AB*	*KW*	AB	KW	—				

Table 1: Overview of data from the subject description of dissertations from selected universities in the Czech Republic in various information systems and selected national or international services [Legend: IR = institutional repository, AB = abstract, KW = keyword, SH = subject heading system]

7 catalogues also use the well–known CONSPECTUS categorisation scheme.

It is gratifying that Czech dissertations have the possibility of being better searchable in terms of subject, even if only through OPACs. Unfortunately, for now only separately in each catalogue. Records of Czech dissertations are not currently placed – unlike in other developed countries – into the **Czech national union catalogue (CASLIN)**. This is a great pity, as we could then have access to the whole collection of dissertation records from a single location. Another theme could be the idea of enriching the records in the repositories of universities with certain data from catalogues. The idea is logical, yet no practical solution can be anticipated at this time.

Dissertations and database of the Czech National Bibliography

In the Czech Republic, the nationwide registration of all defended dissertations was conducted in the past. The traditional system disappeared after the cancellation of the applicable legislation (after 1990). It is a pity that this registration could not continue, as similar systems successfully operate in other countries.

As part of the survey of the processing of dissertation subjects in catalogues, the database of the Czech National Bibliography (CNB, **http://aleph.nkp.cz/F/**) was also checked. It is true that the CNB today registers a number of dissertation autoabstracts, as they are officially published at some universities[8] (with ISBN identifiers). The CNB system gets them as legal deposit. They are subject to subsequent high quality processing in the CNB database,

---

[8] For example, Tomas Bata University in Zlín, Brno University of Technology, VŠB - Technical University of Ostrava, etc.

including detailed subject and systematic description [9]. For example, the record of a dissertation autoabstract by author Petr Maršálek entitled *"Únavové zkoušky ozubených kol = Gear fatigue tests"* (ISBN 978-80-248-2991-3) obtained up to 4 UDC numbers and, in addition to formal subject headings, also two subject headings from the CZENAS[10] controlled vocabulary. Thanks to cooperation between the National Library of the Czech Republic and OCLC, the record is already also in the WorldCat database (**http://www.worldcat.org/oclc/855464854**) and is also currently available on the web in the linked open data structure (in the schema.org and RDFa standards). The record of the autoabstract of the quoted dissertation as a work is freely available, and has a URI identifier: **http://worldcat.org/entity/work/id/1374607131**. The record of this work will be linked to the records of Czech controlled subject headings including their English equivalents (see the block "about" in Figure 3), as soon as the National Library of the Czech Republic issues them as linked data with the relevant URI identifiers. Currently, when you try to click on it the following English message appears: *"This is a placeholder reference for a Topic entity, related to a WorldCat Entity. Over time, these references will be replaced with persistent URIs to... and other Linked Data resources"*.

Figure 3: Part of the record of a Czech dissertation autoabstract as a work from an experimental OCLC linked open data server

We can currently only regret that Czech dissertations are not processed in the system of national bibliography. High quality records of them could form part of important international systems, including the Google Books system, which takes records from the WorldCat catalogue. This would improve their searchability several fold, including from the mobile devices of end users.

---

[9] The cataloguing of published dissertation autoabstracts is also performed at the level of university catalogues.

[10] A complete record of the autoabstract of the given dissertation is in the CNB database, available at URL: **http://aleph.nkp.cz/F/?func=direct&doc_number=002466242&local_base=CNB**.

Searching for doctoral dissertations by subject abroad

In the area of processing and searching for dissertations by subject, the situation abroad is substantially better than that in the Czech Republic. In many countries, the operation of various systems at both national and supranational level for their processing and online searching from a single location, including access to their texts, is a matter of course. The processing of subjects using controlled vocabularies of the subject headings schemes or classification schemes, which are in many cases already issued free of charge in the form of linked open data, is commonplace. Their searchability is undoubtedly at a high level, and also through the Google search engine. The situation in the United States and in Germany is concisely characterised below as an example.

United States

The United States (U.S.) has two systems today, and their databases have a similar number of dissertation records (Procious, 2014, p. 144). All doctoral dissertations are registered.

The first is the **PQDT (ProQuest Dissertation & Theses)** database operated by ProQuest. It holds around 3,500,000 dissertation records and also master's theses. All doctoral dissertations from the United States are registered (around 40,000 doctoral dissertations are defended annually). Around 700 universities are involved. The bibliographic records in the PQDT database are very detailed as regards their formal description, while ISBN identifiers are also assigned. Online searches for dissertations are provided by ProQuest in the FAST search system (a new one is under preparation). As regards subjects, the record is furnished with an author abstract, while ProQuest professionals classify dissertations into their own universal categorisation scheme (ProQuest Subject Categories). Lexical equivalents of the categorisation are in addition also localised in the subject headings field. The record also contains author keywords as a rule. Thanks to the cooperation with universities, it is possible to transfer records from the PQDT database (selected useful data) into university catalogues, where they are subject to high quality processing, in particular as regards the dissertation subject. Subsequently the records are transferred to the WorldCat database (OCLC). The searchability of dissertations by subject is further increased in some professional bibliographic databases (PsycINFO, ERIC etc.), into which ProQuest transfers its records by agreement. In them, the records are further supplemented with more detailed subject terms from thesauri and numbers from special classification schemes.

The second database in the United States that contains a large number of dissertation records is the **WorldCat** database. The records are sourced from university catalogues, are furnished with Library of Congress Subject Headings (LCSH) and automatically also headings from the new FAST system (OCLC) and also LCC or DDC classification numbers. Thanks to the technological maturity of the OCLC, the records of these dissertations are today issued in the form of linked open data, while the linking to subject heading systems is actually live. For example, a dissertation by author N. R. Kale entitled *"A case study on robustness of Dynamic Time Warping for activity recognition using wearable computers"* (URI: **http://www.worldcat.org/oclc/846506856**), defended at the University of Texas in Dallas in 2012 has, in addition to the abstract, up to 8 subject headings linked from the record of the work (the URI of the dissertation in question as a work is **http://worldcat.org/entity/work/id/1356254544**) to the records of relevant subject heading

systems available on the web: LCSH (for example the heading "Dynamic programming", URI: **http://id.loc.gov/authorities/subjects/sh85040313**) and FAST (for example the heading "Wireless sensor nodes", URI: **http://id.worldcat.org/fast/1750044**).

The searchability of dissertations from the United States by subject is thus at a high level, even under the conditions of a growing semantic web.

Germany

Searchability by subject for all dissertations defended at German universities is similarly favourable. Thanks to a law on legal deposit of dissertations, they are now completely registered in the German National Bibliography system. The records are accessible from the catalogue of the German National Library, meaning also from a single location. The publication of the records excels through its precision – all physical records of essential "things" (objects) already have unique URI identifiers, including records of subject headings and name authority data. The dissertation subject is represented primarily through the controlled Subject Headings Authority File (SWD – Schlagwortnormdatei), which is today part of the Integrated Authority File (GND – Gemeinsame Normdatei), also already issued in the linked open data structure. An example can be a dissertation by author Florian Klein entitled *"Metadaten-Verwaltung in einem verteilten RAM-basierten Speicherdienst"*, defended at the University of Düsseldorf in 2015 (**http://d-nb.info/1079652574**). The dissertation is assigned a URN identifier (access to the text is possible from two locations), the record is primarily linked to the national authority record of the author (**http://d-nb.info/gnd/187515727**), the records of two referees (**http://d-nb.info/gnd/1018114750** and **http://d-nb.info/gnd/122535316**) and a total of 6 records of controlled subject headings (in German): Speicher <Informatik> (**http://d-nb.info/gnd/4077653-0**), Metadaten (**http://d-nb.info/gnd/4410512-5**), Verwaltung (**http://d-nb.info/gnd/4063317-2**), Dienst <Informatik> (**http://d-nb.info/gnd/4835035-7**), Speicherverwaltung (**http://d-nb.info/gnd/4182146-4**), Anwendungssoftware **http://d-nb.info/gnd/4120906**). In addition, there is also the classification number of the German National Library classification scheme. See Figure 4.
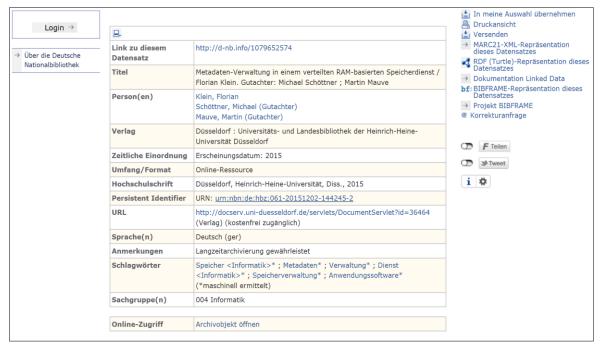
Figure 4: Record of a German dissertation in the catalogue of the German National Library with a URN identifier and also with hyperlinks to GND (Gemeinsame Normdatei) subject heading records.

Thanks to cooperation with OCLC, records of German dissertations are also available in the WorldCat database, and are therefore also available in the open linked data form. Clicking on German national subject headings still results in the message quoted above in the text for the example description of a single Czech dissertation autoabstract. In fact, this message is no longer completely true, as German subject headings already have URIs, and it is anticipated that they will soon be correctly linked through the OCLC experiment.

Conclusion

The digital repositories at Czech universities, as well as the two national services that register and make dissertations available, are developing successfully and growing in terms of data volumes. The records of dissertations from some universities are also already transferring to international services (BASE, OpenAIRE and DART-Europe – see Table 1). Yet they are still lacking something fundamental, namely better and higher quality access to searches for dissertations by subject. With the growing volumes of records and dissertations themselves, there is the danger of serious losses in terms of quality and effective searching in the future. Just finding "something" from a specific quantity of content–relevant dissertations may possibly satisfy the layman (student etc.) at a given moment, yet will not satisfy a professional searcher, whose task it might be to search for all dissertations on a specific subject. A small survey in the Czech Republic showed that there is some hope for improved description of subjects, and this through the catalogues of university libraries. Many libraries (but not all) allocate subject headings, UDC classification numbers and the Conspectus category system. There is a chance that in the future dissertation records may be linked on the web to records of these subject heading systems and classifications. The exposure of some knowledge organization systems (KOS) in the linked open data structure is already a reality, while others are under preparation. This is a powerful technology that could help change the unfavourable Czech situation in the area of dissertation subject descriptions. Can we, for example, hope that it will be possible – at least in some way – to integrate the descriptions of dissertation subjects in

repositories and catalogues separately? Can we hope that records of Czech dissertations will also be entered into the Czech national union catalogue CASLIN? Yet these are questions for a different discussion.

References

CHARVÁTOVÁ, Michaela, 2016. *Sjednocování věcného popisu agregovaných záznamů v repozitáři NUŠL*. Praha. 121 p., 2 suppl. Diplomová práce (Mgr.). Univerzita Karlova, Filozofická fakulta. Available from: **https://is.cuni.cz/webapps/zzp/detail/161853/**.

McCUTCHEON, Sevim aj., 2008. Morphing Metadata: Maximizing Access to Electronic Theses and Dissertations. *Library Hi Tech*. **26**(1), 41-57. Also available commercially from (DOI): **http://dx.doi.org/10.1108/07378830810857799**.

MIDDLETON, Cedar C. aj., 2015. A Process for the Original Cataloging of Theses and Dissertations. *Cataloging & Classification Quarterly*. **53**(2), 234-246. Also available commercially from (DOI): **http://dx.doi.org/10.1080/01639374.2014.971997**.

PROCIOUS, Aaron W., 2014. WorldCat, the Other ETD Database: An Exploratory Study. *The Reference Librarian*. **55**(2), 144-150. Also available commercially from (DOI): **http://dx.doi.org/10.1080/02763877.2014.880276**.